

FIG. 1



Similarity of Image T_1 and T_2

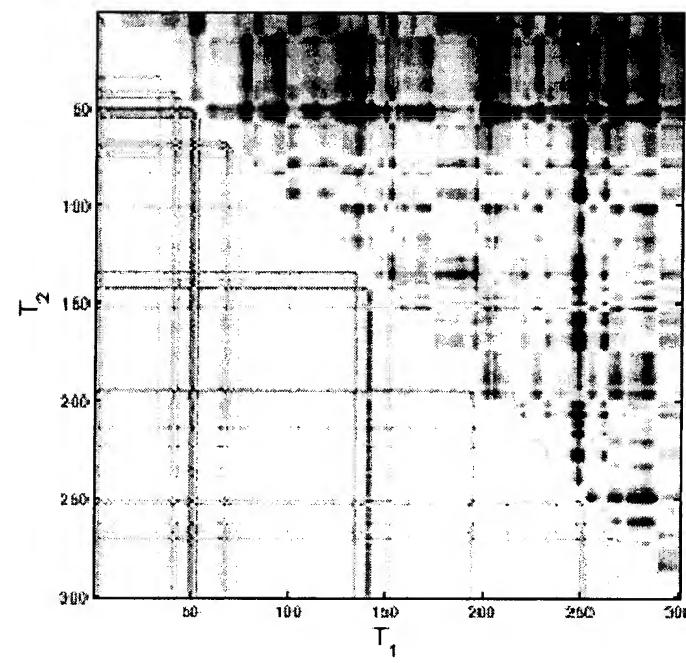


FIG. 2

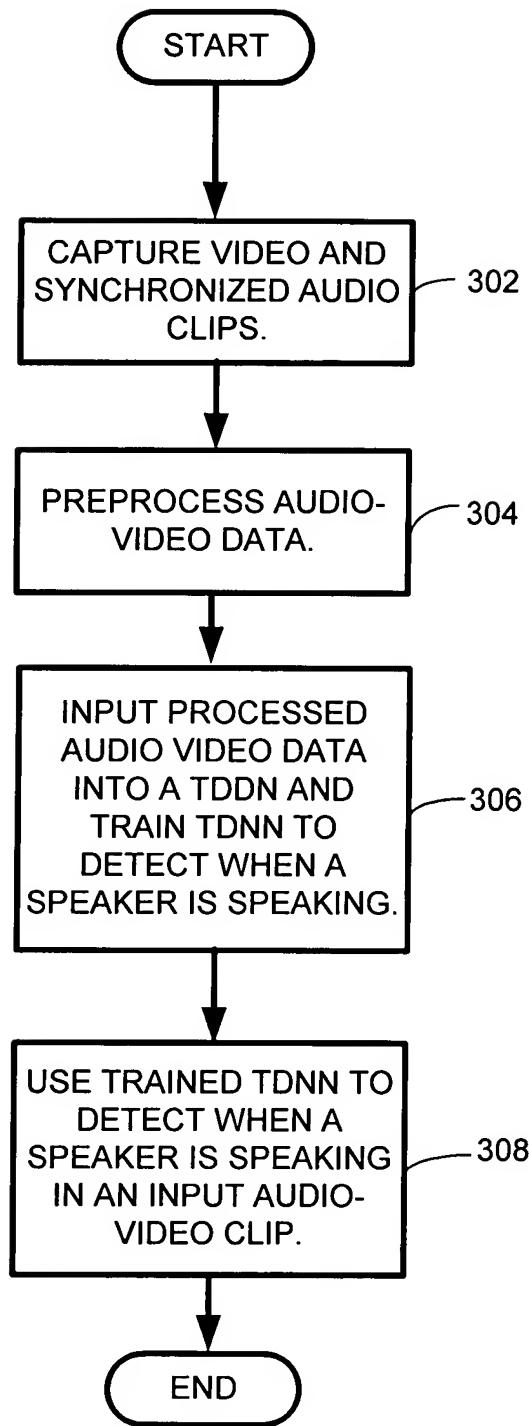


FIG. 3

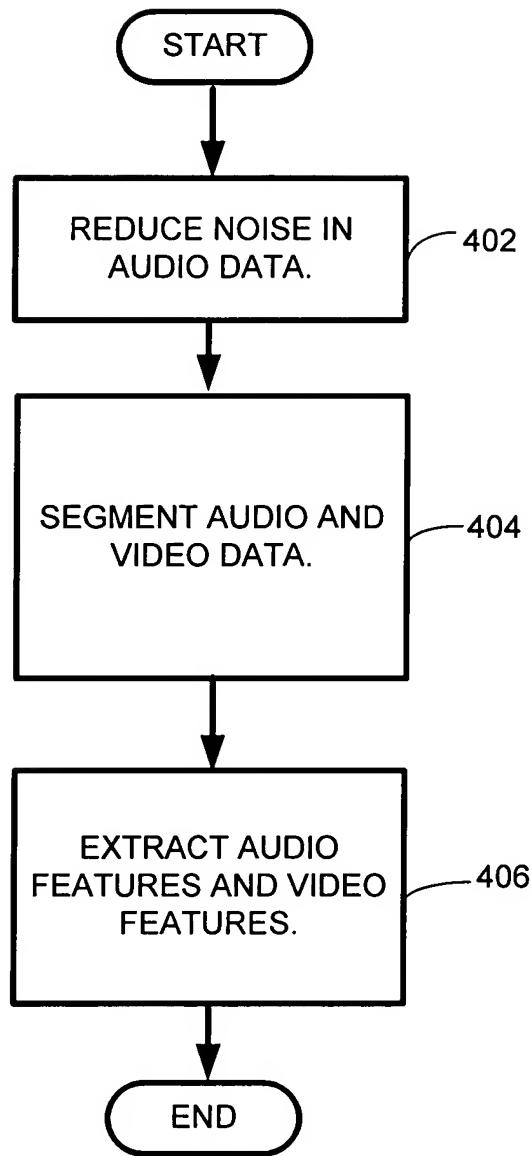
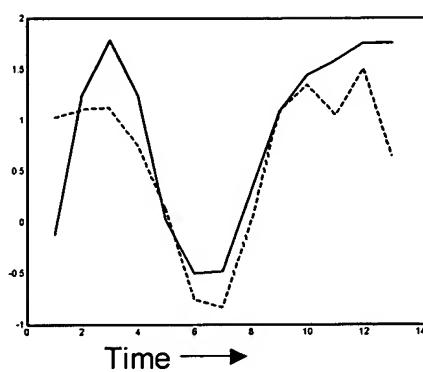


FIG. 4



Mouth
Openness

— Audio
Energy

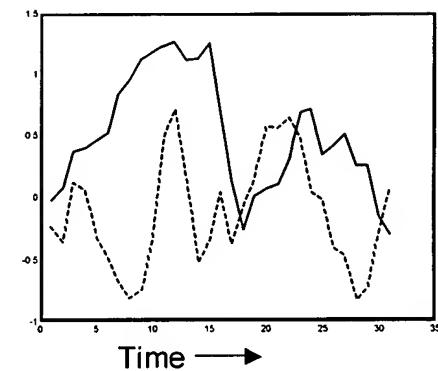
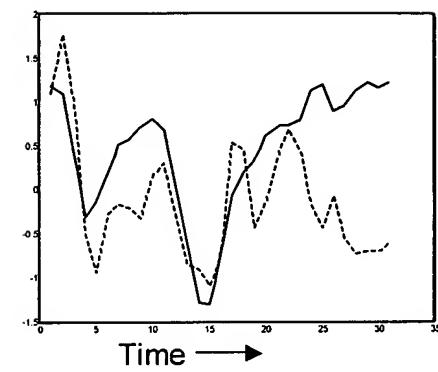
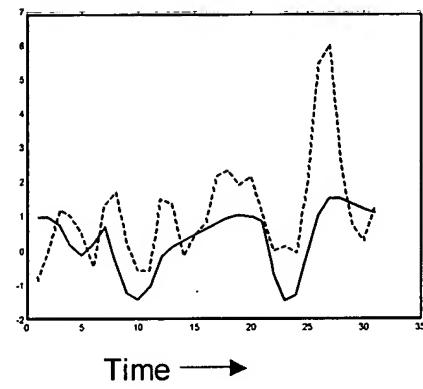
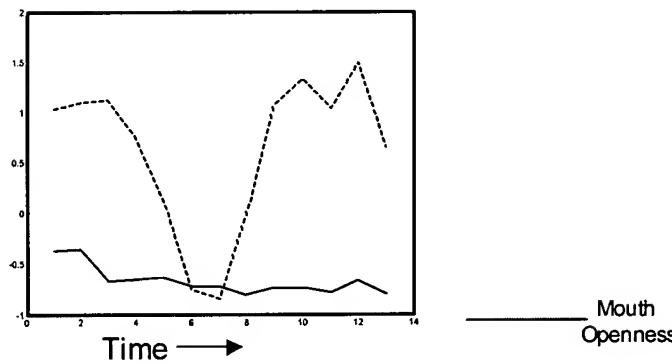


FIG. 5A



Mouth
Openness

— Audio
Energy

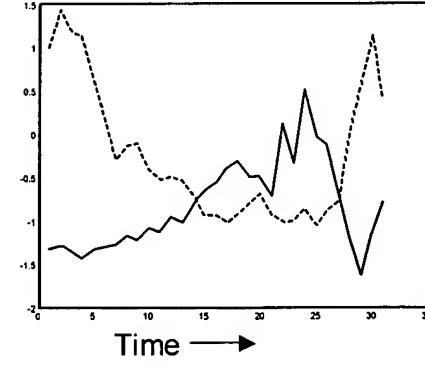
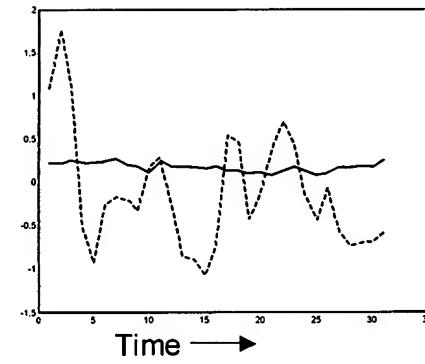
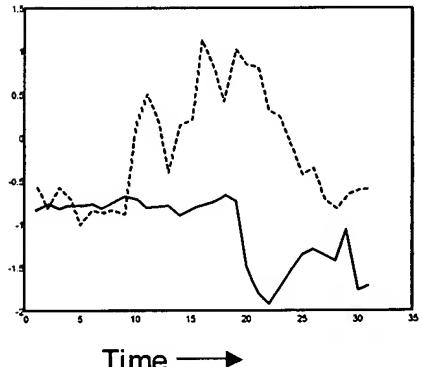


FIG. 5B

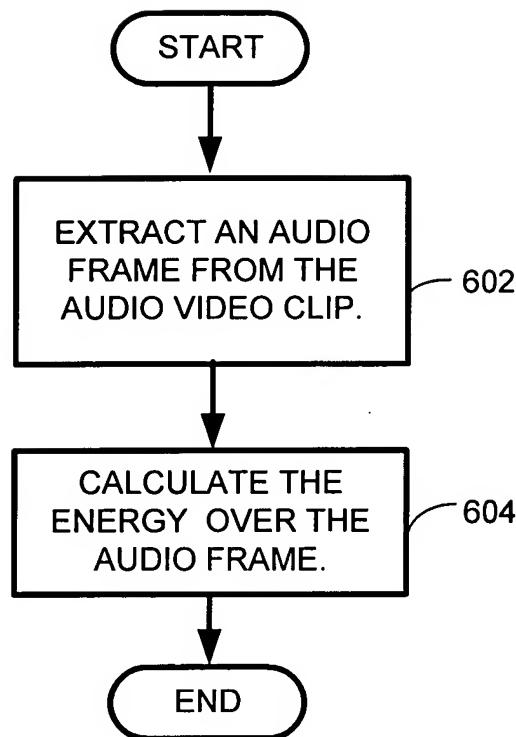


FIG. 6



FIG. 8

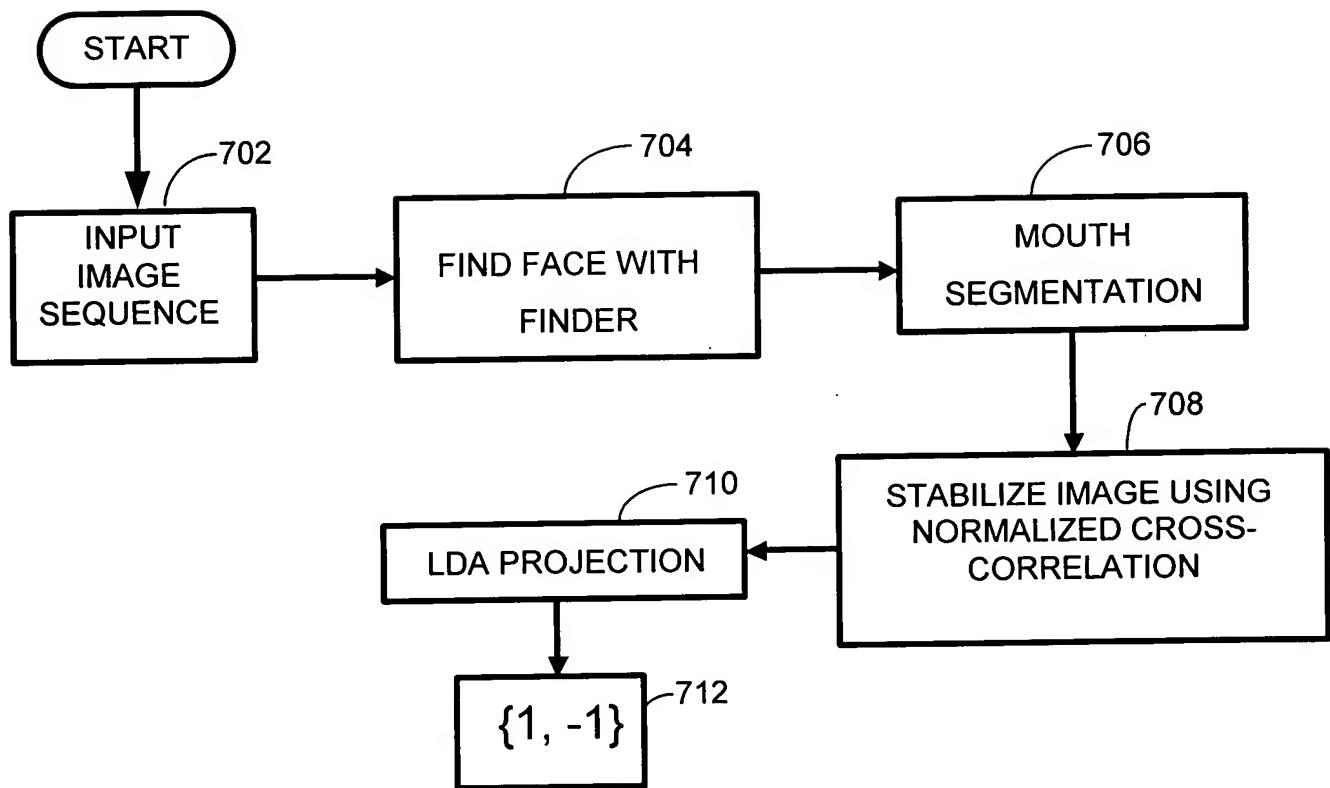


FIG. 7

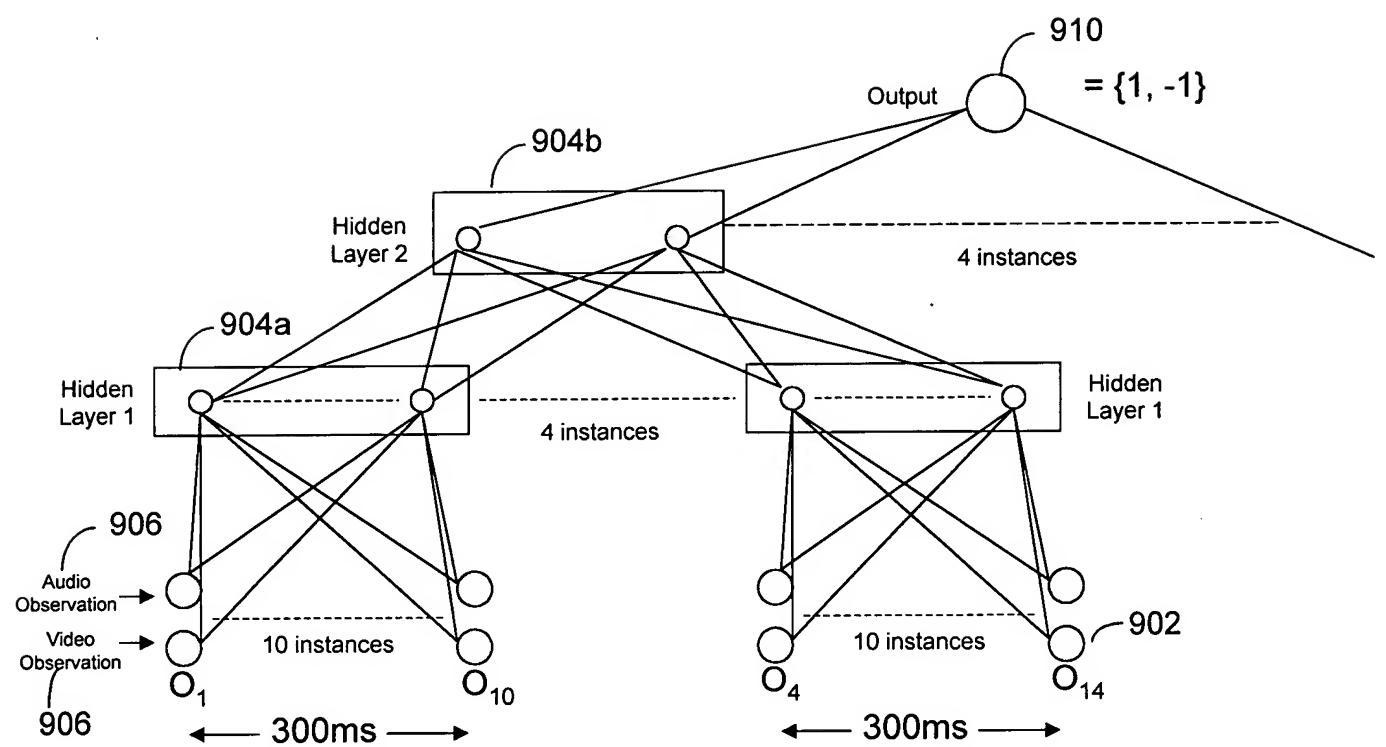


FIG. 9

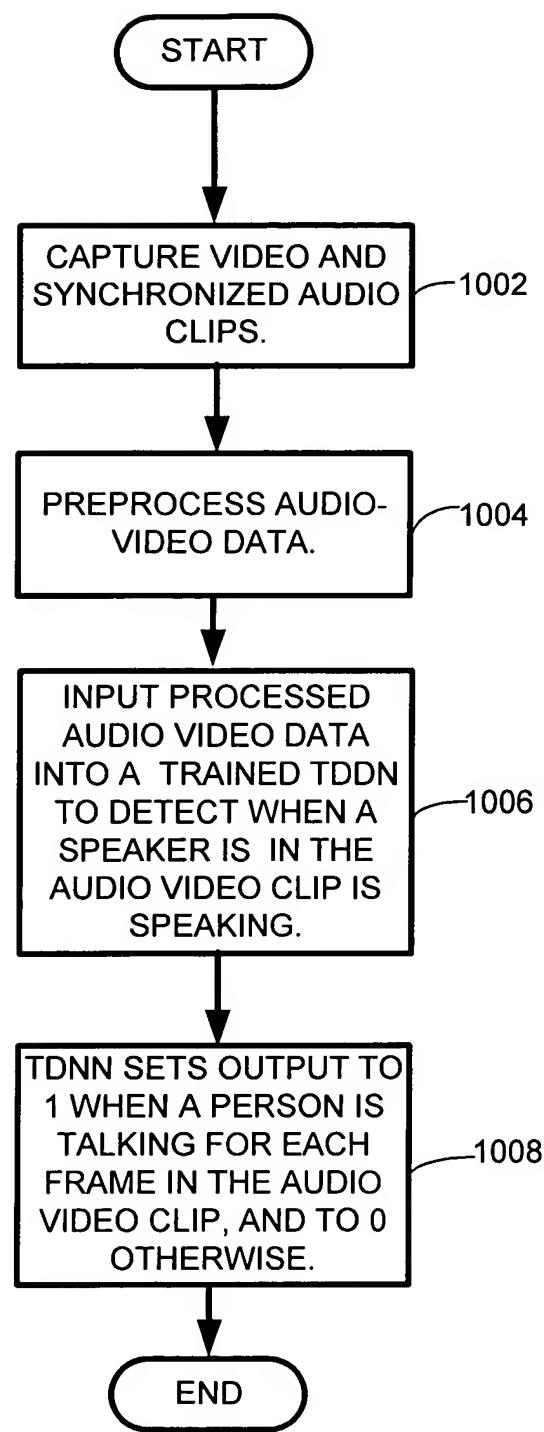


FIG. 10



		Detected		
		Ross Speaking	Ashish Speaking	Silence
Actual	Ross Speaking	7	0	0
	Ashish Speaking	0	18	2
	Silence	0	4	1

Video Only: Accuracy = 81.25%

FIG. 11A

		Detected		
		Ross Speaking	Ashish Speaking	Silence
Actual	Ross Speaking	3	0	4
	Ashish Speaking	2	2	16
	Silence	2	0	3

Audio Only: Accuracy = 25%

FIG. 11B

		Detected		
		Ross Speaking	Ashish Speaking	Silence
Actual	Ross Speaking	7	0	0
	Ashish Speaking	0	20	0
	Silence	0	3	2

Audio & Video: Accuracy = 90.625%

FIG. 11C